

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

OncoBalance Model: A Systems-Based Probabilistic Framework for Cancer Risk and Progression

Abstract

Cancer emerges from a progressive imbalance between genomic damage, cellular repair capacity, and immune surveillance. The OncoBalance Model is proposed as a quantitative systems framework that integrates tumor mutational burden (TMB), oncogenic gene-expression activity, epigenetic disruption, DNA repair capacity, and immune response into a unified probabilistic score of malignant transition. Unlike deterministic threshold models, this approach defines cancer as a continuous probabilistic transition process rather than a binary event. The framework is compatible with large-scale genomic datasets and is intended for research-stage predictive modeling, biological stratification, and future translational development. [file:150]

1. Introduction

Cancer is not a binary event but a dynamic systems failure. Existing approaches often rely on discrete biomarkers yet lack a unified mathematical structure that expresses how genomic instability, failed repair, and impaired immune surveillance interact within a single probabilistic system. The central hypothesis of the OncoBalance Model is that malignant transition emerges when the balance between oncogenic burden and regulatory protection shifts beyond a critical tissue-dependent regime. [file:150]

By quantifying this imbalance, the model seeks to support a transition from reactive diagnostics toward proactive and systems-based risk stratification. At this stage, the framework should be interpreted as a research-oriented probabilistic model whose biological coherence is strong at the conceptual level, while its clinical validity remains dependent on formal calibration and prospective validation. [file:150]

2. Model Definition

2.1 Oncogenic Information Function

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
 No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
 See full “License and Conditions of Use” notice for details.
 In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
 Author: Carlos Daniel Borrego Romero

The oncogenic information content H_q represents the accumulated molecular burden or destabilizing biological signal within the cellular system:

$$H_q = w_1 \cdot \text{TMB} + w_2 \cdot E_{\text{gene}} + w_3 \cdot E_{\text{epi}}$$

where:

- **TMB** is tumor mutational burden.
- E_{gene} is an oncogenic gene-expression score.
- E_{epi} is an epigenetic disruption score, for example global hypomethylation or another validated epigenetic-instability proxy.
- w_1, w_2, w_3 are learned non-negative weights to be optimized during model training.

This formulation is mathematically valid as a weighted additive structure. However, because the three components are not naturally measured on the same scale, a practical implementation must specify normalization, directionality, and transformation rules before aggregation. In particular, all three inputs should be harmonized to comparable scales so that the learned weights reflect biological relevance rather than arbitrary differences in units. [file:150]

2.2 Biological Balance Equation

Cancer-related imbalance is modeled as the ratio between oncogenic burden and the system's protective mechanisms:

$$B = \frac{H_q}{R_{\text{repair}} + I_{\text{immune}} + \epsilon}$$

where:

- R_{repair} denotes DNA repair capacity.
- I_{immune} denotes effective immune-system activity.
- $\epsilon > 0$ is a small stabilization constant introduced to prevent division by zero under conditions of near-complete regulatory collapse.

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

This ratio is internally coherent: increases in oncogenic burden raise B , while increases in repair capacity or immune protection lower B . The value of ϵ should not be treated as purely arbitrary; although $\epsilon = 0.1$ is a workable initial choice, it should be considered a tunable hyperparameter subject to sensitivity analysis because it affects model behavior in extreme low-protection states. [file:150]

2.3 Cancer Probability Function

The probability-like score of malignant transition is mapped through a logistic function:

$$P(\text{cancer}) = \frac{1}{1 + e^{-k(B-\theta)}}$$

where:

- k is the slope or sensitivity parameter.
- θ is a tissue-specific critical threshold.

This function maps the imbalance term into the interval (0,1). As a mathematical mapping, it is appropriate and ensures bounded output. However, the resulting value should initially be described as a **probability score** or **malignancy-risk score** rather than as a clinically calibrated probability unless calibration has been formally demonstrated on external datasets. [file:150]

2.4 Mathematical Coherence and Boundary Behavior

The model behaves coherently under limiting conditions. If H_q approaches zero while repair and immune protection remain high, then B approaches zero and the logistic score approaches a low-risk regime. If oncogenic burden becomes large while protection is weak, then B increases substantially and the logistic score approaches one. If repair and immune protection collapse simultaneously, the denominator remains finite because of ϵ , preventing singularities and preserving numerical stability. [file:150]

The model also behaves intuitively in intermediate states. High TMB with strong immune activity can produce moderated imbalance, while low TMB with defective repair may still generate non-trivial risk if protection is insufficient. This is one of the conceptual strengths of the framework because it allows different oncogenic routes to converge on a common systems-level risk representation. [file:150]

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

2.5 Operationalization of Model Components

For the model to be empirically implementable and clinically interpretable, each theoretical component must be mapped onto observable dimensions, candidate indicators, and explicit scaling rules. At this stage, operational definitions remain framework-level and can be adapted to specific tumor types, assay platforms, and cohorts. [file:150]

2.5.1 Oncogenic burden H_q

The latent oncogenic information term H_q is constructed from three observable domains:

- **Tumor mutational burden (TMB).** Primary indicators include nonsynonymous mutations per megabase from whole-exome sequencing or validated targeted gene panels, as well as ctDNA-based TMB estimates for longitudinal monitoring. [file:150][web:141]
- **Oncogenic gene-expression activity (E_{gene}).** Primary indicators include tumor-type-specific transcriptomic signatures capturing oncogenic programs such as proliferation, MYC activation, epithelial-mesenchymal transition, or cell-cycle dysregulation, derived from bulk RNA sequencing or equivalent expression platforms. [file:150][web:137]
- **Epigenetic disruption (E_{epi}).** Primary indicators include global hypomethylation indices, cancer-associated methylation classifiers, or other validated epigenetic-instability scores obtained from tissue or liquid-biopsy methylation assays. [file:150][web:145]

All three inputs should be transformed to a common scale, for example 0–1 within tumor type, before weighted aggregation. This ensures that the learned weights (w_1, w_2, w_3) reflect biological importance rather than raw-unit differences. [file:150]

2.5.2 DNA repair capacity R_{repair}

The repair-capacity term R_{repair} represents the system's ability to correct genomic damage. Candidate indicators include homologous-recombination deficiency composite scores such as LOH, TAI, and LST; mismatch-repair deficiency and MSI status measured by immunohistochemistry, PCR, or NGS; and pathway-level expression signatures of BER, NER, HR, and related repair pathways. [file:150][web:140][web:146]

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

These indicators can be combined into a normalized composite repair score in which higher values correspond to stronger repair capacity. When necessary, raw measures such as HRD scores, which increase with deficiency, should be directionally aligned so that higher R_{repair} always reflects greater protection. [file:150]

2.5.3 Immune activity I_{immune}

The immune-activity term I_{immune} quantifies effective anti-tumor immune surveillance. Candidate indicators include densities of cytotoxic CD8+ T cells and NK cells measured by immunohistochemistry or inferred from deconvolved bulk RNA; proportions of CD8+ cells relative to immunosuppressive populations such as Treg or MDSC, forming an Effective Immunity Index; and cytolytic activity scores or IFN- γ -related inflammatory signatures. [file:150][web:137][web:138]

These features can be integrated into a normalized immune-activity index on a 0–1 scale, where higher values indicate stronger effective immune surveillance. [file:150]

2.5.4 Imbalance velocity ΔB

The imbalance-velocity term ΔB captures the rate of change in the balance index over time. Observable inputs are serial measurements of the same components used to construct B , including longitudinal ctDNA TMB, repeated methylation signatures, serial transcriptional or protein-based immune markers, and related liquid-biopsy measurements. [file:150][web:139][web:145]

To ensure that ΔB reflects biological evolution rather than measurement noise, future implementations should specify minimum intervals between measurements, assay comparability across time points, and smoothing rules for noisy trajectories. [file:150]

2.5.5 Scaling and normalization rules

Because TMB, expression scores, methylation indices, repair metrics, and immune metrics operate on heterogeneous numerical scales, the following normalization principles are recommended:

- Within each tumor type and cohort, primary indicators should be transformed to comparable scales, such as min–max or percentile-based scaling to 0–1 with explicit directionality, or z-score standardization followed by bounded transformation when needed for stability. [file:150][web:141][web:145]

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.

No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.

See full “License and Conditions of Use” notice for details.

In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.

Author: Carlos Daniel Borrego Romero

- Inputs to H_q , R_{repair} , and I_{immune} should be normalized before aggregation, and weights w_i should remain non-negative to preserve interpretability. [file:150]
- The stabilization constant ϵ should be chosen relative to the normalized scale of $R_{\text{repair}} + I_{\text{immune}}$, for example in the 0.01–0.1 range when both terms are bounded in $[0,1]$, and should be examined in sensitivity analyses. [file:150]
- The resulting balance index B can then be compared consistently across patients within a given tumor type and passed through the logistic mapping to yield a malignancy-risk score. [file:150]

3. Methods

3.1 Data Sources and Integration

The framework is designed to interface with high-dimensional multi-omic repositories for both retrospective model development and future real-time inference. Static training data may be drawn from large-scale cancer-genomics resources, while dynamic inference may incorporate liquid-biopsy-derived biomarkers such as circulating tumor DNA for TMB monitoring, plasma methylation signals for epigenetic disruption, and cytokine or immune signatures for immune-state estimation. [file:150]

For methodological rigor, future versions of the model should distinguish clearly between **discovery datasets**, **external retrospective validation cohorts**, and **prospective longitudinal cohorts**. This separation is essential in medical modeling because retrospective fit alone is insufficient to justify clinical interpretability or transportability. [file:150]

3.2 Variable Construction and Proxy Mapping

Theoretical variables must be linked to reproducible operational proxies.

- **Repair Capacity** (R_{repair}) may be estimated from pathway-expression signatures, homologous-recombination deficiency measures, mismatch-repair deficiency markers, genomic-scar assays, or tissue-specific composite repair scores. [file:150]
- **Immune Activity** (I_{immune}) may be estimated through an Effective Immunity Index based on cytolytic CD8+ activity relative to immunosuppressive populations such as Treg or MDSC. [file:150]

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

At this stage, the manuscript should explicitly state whether these proxies are derived from bulk RNA sequencing, deconvolution methods, histopathology, blood-based biomarkers, or integrated multi-modal assays. This clarification is necessary because different measurement strategies may not be interchangeable. [file:150]

3.3 Model Training and Optimization

The weights $\mathbf{w}$ and parameters (k, θ) are proposed to be optimized using an Elastic Net regularized supervised-learning protocol. This is a reasonable choice for high-dimensional biomedical data because the combined L_1 and L_2 penalties can reduce instability under correlated predictors while retaining feature selection capacity. [file:150]

For a medically robust implementation, the training procedure should further specify patient-level data partitioning, prevention of data leakage, tissue-stratified evaluation, missing-data handling, and whether parameter estimation is global or tissue-specific. These details are critical because apparent mathematical performance can otherwise be inflated by methodological leakage rather than real biological generalization. [file:150]

3.4 Explainability

SHAP analysis is proposed to decompose the contribution of each variable. This is appropriate as an explainability layer, especially in a translational setting where clinicians and researchers need to understand whether a high-risk score is being driven primarily by mutational burden, repair failure, epigenetic disruption, or immune collapse. [file:150]

However, explainability should be presented as a transparency mechanism rather than as evidence of validity. A clinically interpretable model is not necessarily a clinically valid model unless its predictions remain stable, calibrated, and externally reproducible. [file:150]

3.5 Sensitivity and Velocity Analysis

A temporal instability term is introduced as

$$\Delta B = \frac{B(t_2) - B(t_1)}{t_2 - t_1}$$

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

which tracks the velocity of imbalance over time. This extension is mathematically valid and biologically useful because many clinically relevant transitions are dynamic rather than static. A rising ΔB may indicate that the system is moving toward a higher-risk regime even before conventional diagnostic thresholds are crossed. [file:150]

Future implementations should specify the minimum time interval between measurements, assay comparability across time points, and smoothing rules for noisy longitudinal data so that ΔB captures biological change rather than measurement fluctuation. [file:150]

4. Validation and Calibration Plan

Because this is a medical and translational model, formal validation must extend beyond mathematical plausibility and include endpoint specification, cohort design, calibration analysis, external reproducibility, and clinical-utility assessment. The purpose of this section is to define the minimum validation architecture required for the OncoBalance framework to evolve from a conceptually coherent model into a scientifically robust predictive system. [file:150]

4.1 Clinical endpoints

A central requirement for medical validation is a precise endpoint definition. Depending on the intended use case, the target outcome may be one of the following:

- current malignant status at the time of sampling,
- short-term progression from premalignant or high-risk state to confirmed malignancy,
- recurrence after treatment,
- molecular residual disease persistence,
- or treatment-response probability under a defined therapeutic context. [file:150]

These endpoints should not be mixed without explicit stratification. A score optimized for current malignancy detection is not automatically valid for recurrence prediction or treatment guidance. [file:150]

4.2 Cohort structure

Validation should proceed in clearly separated stages:

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.

No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.

See full “License and Conditions of Use” notice for details.

In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.

Author: Carlos Daniel Borrego Romero

1. **Discovery/training cohorts** for feature engineering, initial parameter learning, and internal cross-validation.
2. **External retrospective cohorts** for transportability testing across institutions, assay platforms, and patient populations.
3. **Prospective longitudinal cohorts** for temporal validation, serial monitoring, calibration, and assessment of clinical utility. [file:150]

Where possible, cohorts should be stratified by tumor type, disease stage, treatment context, and assay platform. Pan-cancer pooling may be useful for discovery, but clinical validity must ultimately be demonstrated within biologically and clinically coherent subgroups. [file:150]

4.3 Internal validation protocol

A medically robust internal validation protocol should include:

- patient-level separation between training, validation, and test sets,
- nested cross-validation for hyperparameter tuning,
- leakage control between molecular profiles and outcome labels,
- predefined missing-data procedures,
- and tissue-stratified performance reporting. [file:150]

These steps are essential because apparent predictive strength can be inflated by overfitting or leakage in high-dimensional multi-omic data. [file:150]

4.4 External validation and reproducibility

The model should be evaluated on at least one independent retrospective cohort and, ideally, a prospective cohort. External validation should examine whether the score retains discrimination, calibration, and biological interpretability when applied outside the discovery dataset. [file:150]

Performance should be reported separately for each tumor type or clinically coherent tumor family. Strong pooled performance is insufficient if calibration fails in specific tissues or stages. [file:150]

4.5 Performance metrics

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

The training process may use log-loss as an optimization objective, but the final evaluation should include multiple complementary metrics:

- AUROC,
- area under the precision-recall curve,
- Brier score,
- calibration slope and intercept,
- calibration plots,
- sensitivity and specificity at clinically justified thresholds,
- and decision-curve analysis for evaluating potential clinical usefulness. [file:150]

Where relevant, subgroup analyses should also be reported by tumor type, stage, age, sex, ancestry, and assay platform. [file:150]

4.6 Calibration of the probability score

The logistic mapping guarantees a bounded probability-like score, but it does not by itself guarantee calibrated clinical risk. For the score to be interpreted as a probability in clinical or translational settings, formal calibration must be demonstrated on held-out and external data. [file:150]

Potential approaches include post-hoc calibration methods such as Platt scaling or isotonic regression, followed by assessment of calibration slope, intercept, and reliability curves. Until such calibration is demonstrated, the output should be described as a malignancy-risk score rather than a clinically validated probability. [file:150]

4.7 Longitudinal validation of ΔB

The dynamic term ΔB requires a separate validation strategy. It should be tested against clinically meaningful temporal outcomes, such as early relapse, molecular residual disease expansion, transition from premalignant states, or progression under surveillance. [file:150]

Its usefulness will depend not only on predictive strength but also on its stability across repeated sampling and assay variability. Therefore, longitudinal validation should report both discrimination and temporal robustness. [file:150]

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

4.8 Incremental value over existing models

Because oncology already uses strong biomarker frameworks, OncoBalance should not only be tested in isolation but also against existing standards. Validation studies should assess whether the model adds predictive or stratification value beyond single biomarkers or established composite signatures. [file:150]

This can be examined through nested-model comparison, net reclassification improvement, calibration improvement, and decision-curve analysis. The most defensible future claim is not merely that the model works, but that it improves joint interpretation beyond current one-dimensional biomarkers. [file:150]

4.9 Bias, fairness, and implementation safeguards

Clinical validation should also examine representation bias across ancestry, sex, age, disease stage, and assay platform. In high-dimensional oncology datasets, a model may appear globally strong while systematically underperforming in underrepresented populations or rare tumor contexts. [file:150]

Accordingly, every major validation phase should include subgroup performance reporting, uncertainty estimation, and explicit acknowledgment of settings where the model should not yet be used. [file:150]

5. Results (Expected Behavior)

The model is expected to reproduce several biologically plausible patterns:

- **High TMB + low repair:** $P(\text{cancer})$ rises toward a high-risk regime.
- **High TMB + strong immune response:** B is moderated, reflecting partially controlled immunogenic tumors.
- **Low TMB + defective repair:** intermediate risk remains possible, capturing tumors driven by repair dysfunction or high-impact driver events rather than bulk mutational accumulation alone. [file:150]

At this stage, these behaviors should be presented as **expected qualitative behaviors** rather than as empirically established results unless validated on independent data. This wording is important in a medical manuscript to avoid overstating model maturity. [file:150]

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

6. Discussion

6.1 Key Contributions

The main contribution of the OncoBalance Model is the unification of disparate biomarkers into a single systems-level imbalance index. Rather than treating isolated mutations or immune markers as independent endpoints, the framework attempts to capture the balance between oncogenic pressure and biological protection. [file:150]

A second contribution is the move away from binary diagnostics toward a graded probabilistic representation. This is conceptually aligned with the reality that malignant transformation is often gradual, heterogeneous, and tissue-dependent rather than governed by a universal deterministic threshold. [file:150]

6.2 Clinical Interpretation and Caution

The model has potential value for risk stratification and longitudinal monitoring, especially in liquid-biopsy-oriented settings. Nevertheless, therapeutic implications should be framed cautiously. A high score driven by low immune protection may support the hypothesis that immune dysfunction is important in that case, while a high score driven by elevated oncogenic burden may indicate a different biological route; however, such distinctions should be considered hypothesis-generating unless linked prospectively to treatment outcomes. [file:150]

Accordingly, direct treatment recommendations should not yet be presented as established consequences of the model. The current evidence supports positioning the framework as a research-stage system for biological interpretation and future translational validation rather than as an autonomous clinical decision tool. [file:150]

7. Limitations

The model requires high-quality, synchronized multi-omic data streams and careful alignment across assay platforms. Tissue-specific calibration of k and θ requires substantial data, ideally including longitudinal follow-up and external validation cohorts. Prospective clinical validation in diverse populations remains pending. [file:150]

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.
No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.
See full “License and Conditions of Use” notice for details.
In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.
Author: Carlos Daniel Borrego Romero

Additional methodological limitations should also be stated explicitly. These include uncertainty in proxy mapping for repair and immune terms, possible instability if inputs are not normalized consistently, cohort imbalance across tumor types, and the risk that a mathematically bounded score may be interpreted clinically as a calibrated probability before such calibration is demonstrated. [file:150]

8. PRAD Mapping Precision Note

As a concrete example of real-data anchoring, the mapping precision of the OncoBalance framework is high when applied to TCGA-PRAD data hosted through the Genomic Data Commons. In this setting, the oncogenic-burden term can be directly represented through harmonized somatic mutation calls for TMB, RNA-seq-based transcriptional profiles for oncogenic expression programs, and HM450K methylation data for epigenetic disruption, while repair-related and immune-related components can be operationally defined from genomic alterations, pathway-level expression signatures, and transcriptome-derived immune infiltration metrics. This makes PRAD an appropriate reference cohort for validating the static structure of H_q , R_{repair} , I_{immune} , B , and the malignancy-risk score. Its main limitation is not the quality of the mapping itself, but the absence of serial liquid-biopsy measurements, which means that PRAD supports rigorous evaluation of the baseline model but not full longitudinal validation of ΔB . [web:184][web:188][web:189][web:190][file:150]

9. Conclusion

The OncoBalance Model provides a mathematically grounded and biologically coherent framework for understanding cancer as a dynamic imbalance system. By integrating genomic damage, repair capacity, and immune surveillance into a unified probabilistic score, it offers a promising structure for predictive modeling and translational oncology research. [file:150]

At the same time, the current version should be described as a **research-stage probabilistic framework** rather than a clinically validated tool. Its next development phase should focus on explicit variable normalization, parameter calibration, external and prospective validation, and careful separation between interpretability, prediction, and clinical utility. [file:150]

LICENSE AND CONDITIONS OF USE.

The content is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license:
<https://creativecommons.org/licenses/by-nc-sa/4.0/>

This model and documentation are provided for non-commercial use only under the CC BY-NC-SA 4.0 license.

No commercial use, profiling, or exploitation of IP logs or trade secrets is permitted.

See full "License and Conditions of Use" notice for details.

In case of conflict or ambiguity, the most restrictive interpretation in favor of non-commercial use, user privacy, and protection of trade secrets shall prevail.

Author: Carlos Daniel Borrego Romero

By using this content and model, you agree to:

Provide appropriate credit, include a link to the above license, and indicate if changes were made.

Not use the material or any derivative works for commercial purposes.

Distribute any adaptations or derivative works only under the same CC BY-NC-SA 4.0 license.

Not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Data logging and trade secrets policy.

The design, implementation, and deployment of this model, as well as any derivatives, do not authorize the collection, storage, or exploitation of IP addresses, unique device identifiers, or other technical traceability data for commercial profiling, targeted advertising, or any other economic exploitation.

Any system implementing this model must limit the logging and retention of IP addresses and other identifiable data strictly to what is necessary to:

(a) maintain technical security and service integrity; and

(b) comply with applicable legal obligations (for example, any minimum log retention required by law).

It is prohibited to use this model or its derivatives to:

capture, store, or exploit trade secrets, proprietary know-how, or confidential information of users or third parties for any purpose other than providing the technical service described in this documentation; or train or improve closed, commercial systems using data that contains confidential information of third parties without their prior, explicit, written consent.

---+EOD+---